



AIML

MODULE PROJECT

Recommendation Systems

TOTAL
SCORE

60

General Instructions:

- 1. Submission of all the parts is expected in 1 notebook only
- 2. Expected submission format: 1 '.ipynb' notebook and 1 '.html' notebook only
- 3. 50% marks will be deducted if insights/steps are missing in the corresponding questions.
- 4. If output for any code cell is missing, 50% marks will be deducted.
- 5. Any kind of Plagiarism will lead to 0 (zero) Marks.

Submission Format:

- 1. '.ipynb' (Jupyter Notebook) and
 - 2. '.html' (Jupyter Notebook > File > Download as > HTML)
- 5 Marks will be deducted if submission in any of the formats is missing.

- **DOMAIN:** Smartphone, Electronics
- **CONTEXT:** India is the second largest market globally for smartphones after China. About 134 million smartphones were sold across India in the year 2017 and is estimated to increase to about 442 million in 2022. India ranked second in the average time spent on mobile web by smartphone users across Asia Pacific. The combination of very high sales volumes and the average smartphone consumer behaviour has made India a very attractive market for foreign vendors. As per Consumer behaviour, 97% of consumers turn to a search engine when they are buying a product vs. 15% who turn to social media. If a seller succeeds to publish smartphones based on user's behaviour/choice at the right place, there are 90% chances that user will enquire for the same. This Case Study is targeted to build a recommendation system based on individual consumer's behaviour or choice.
- **DATA DESCRIPTION:**
 - **author** : name of the person who gave the rating
 - **country** : country the person who gave the rating belongs to
 - **data** : date of the rating
 - **domain**: website from which the rating was taken from
 - **extract**: rating content
 - **language**: language in which the rating was given
 - **product**: name of the product/mobile phone for which the rating was given
 - **score**: average rating for the phone
 - **score_max**: highest rating given for the phone
 - **source**: source from where the rating was taken
- **PROJECT OBJECTIVE:** We will build a recommendation system using popularity based and collaborative filtering methods to recommend mobile phones to a user which are most popular and personalised respectively.

Steps and tasks: [Total Score: 60 points]

1. Import the necessary libraries and read the provided CSVs as a data frame and perform the below steps. [15 Marks]
 - A. Merge all the provided CSVs into one data-frame. [2 Marks]
 - B. Explore, understand the Data and share at least 2 observations. [2 Marks]
 - C. Round off scores to the nearest integers. [3 Marks]
 - D. Check for missing values. Impute the missing values, if any. [2 Marks]
 - E. Check for duplicate values and remove them, if any. [2 Marks]
 - F. Keep only 1 Million data samples. Use random state=612. [2 Marks]
 - G. Drop irrelevant features. Keep features like Author, Product, and Score. [2 Marks]
2. Answer the following questions. [10 Marks]
 - A. Identify the most rated features. [3 Marks]
 - B. Identify the users with most number of reviews. [3 Marks]
 - C. Select the data with products having more than 50 ratings and users who have given more than 50 ratings. Report the shape of the final dataset. [4 Marks]
3. Build a popularity based model and recommend top 5 mobile phones. [5 Marks]
4. Build a collaborative filtering model using SVD. You can use SVD from surprise or build it from scratch(Note: Incase you're building it from scratch you can limit your data points to 5000 samples if you face memory issues). Build a collaborative filtering model using kNNWithMeans from surprise. You can try both user-based and item-based model. [10 Marks]
5. Evaluate the collaborative model. Print RMSE value. [2 Marks]
6. Predict score (average rating) for test users. [2 Marks]
7. Report your findings and inferences. [2 Marks]
8. Try and recommend top 5 products for test users. [5 Marks]
9. Try other techniques (Example: cross validation) to get better results. [3 Marks]
10. In what business scenario you should use popularity based Recommendation Systems ? [2 Marks]
11. In what business scenario you should use CF based Recommendation Systems ? [2 Marks]
12. What other possible methods can you think of which can further improve the recommendation for different users ? [2 Marks]